Postal Regulatory Commission
Submitted 7/8/2015 3:13:40 PM
Filing ID: 92726
Accepted 7/8/2015

Analysis of the Supplemental Report of Dr. Kevin Neels On Behalf of United Parcel Service

Prof. Michael D. Bradley
Department of Economics
George Washington University
Washington D.C, 20052

Introduction

Dr. Kevin Neels submitted a second report, on behalf of United Parcel Service, in Docket No. RM2015-7. His supplementary report both reviews some of the work submitted in his first report and presents a new analysis of city carrier street time, making use of the national Form 3999 data submitted by the Postal Service.[1] The Postal Service has asked me to review Dr. Neels' second report and to provide an evaluation of the potential contributions provided by Dr. Neels' supplemental analysis.

The primary inference that can be drawn from Dr. Neels' supplemental analysis is confirmation of the difficulty of estimating accurate and robust city carrier street time variabilities for parcel volumes. Both the Postal Service and UPS face a challenge in reliably estimating those variabilities, for two reasons. First, despite their growth, parcels still represent only a small proportion of city carrier volumes and times. A few years ago, the time to deliver parcels was only five percent of street time and, although that proportion may have nearly doubled in recent years, it is still a small percentage of total street time.[2] Second, as both UPS and the Postal Service agree, the Postal Service's operational data systems do not include comprehensive and accurate measures of parcels delivered by city carriers.

The key issue, then, is how to best address the challenge of these missing parcel data? The Postal Service's approach was to attempt to rectify the data omission by funding a special study that collected recent information on both parcel volumes and the

[1] See, 'Supplemental Report of Kevin Neels on Behalf of United Parcel Service," Docket No. RM2015-7

[2] See, Report on the City Carrier Street Time Study, Docket No. RM2015-7 at 18.

time spent delivering them. In contrast, UPS's approach was to, essentially, construct pseudo-data to be used in place of actual volume data. Dr. Neels has already provided his concerns about the Postal Service approach, concerns which focus primarily on the amount and quality of the data collected.[3] In the next section, I detail the serious drawbacks of the UPS approach, which include both econometric and implementation problems.

Before addressing those drawbacks, I would like to commend Dr. Neels for resolving three of the five issues raised by his initial report.[4] This reduces the amount of controversy between UPS and the Postal Service, and allows for more focused and serious consideration of the remaining two issues. Dr. Neels now supports the use of the quadratic functional form for estimating city carrier variability equations, he estimates his proposed variability equations on individual, rather than aggregated, volume measures and his modified Proposal 13 model is estimated on all the collected data rather than on average ZIP Code values.[5] This means that the two issues remaining for debate are whether a regular delivery time or a total street time equation should be estimated and whether packages should be included in the regular delivery equation. I will address these two questions in my review of Dr. Neels' two proposed approaches.

---

[3] See, "Report of Kevin Neels on Behalf of United Parcel Service," Docket No. RM2015-7, at 28

[4] Dr. Neels also found a small computational error in the Postal Service's implementation of its deviation parcel equation. I have reviewed Dr. Neels' finding and agree that his proposed correction should be implemented.

[5] This latter issue is not relevant for Dr. Neel's proposed Form 3999 model as it is estimated on a cross-sectional style dataset.

Evaluation of Dr. Neels' Model Based upon the National Form 3999 Dataset

Dr. Neels attempts to use (and augment) the Form 3999 dataset provided by the Postal Service in order to estimate an aggregate street time equation. The appeal of this approach is that there are over 10,000 ZIP Code observations that can be constructed from the Form 3999 dataset. Such a large dataset holds potential for estimating even relatively complex equations.

The primary drawback of this approach is the serious omission of four key volume variables from the dataset. The Form 3999 dataset contains no information on the volumes of mail collected from customers' receptacles, in-receptacle parcels, deviation parcels, and accountables. Because these are all important cost drivers of city carrier street time costs, omitting them would seriously undermine the calculation of attributable city carrier street time costs.

Dr. Neels attempts to address this deficiency by constructing artificial data points for the missing observations. In other words, he attempts to impute values for each observation, for each of the missing variables.[6] Imputation is a well-known technique in econometrics, but typically it is used to fill in a limited number of missing values for a specific variable, usually based upon other information in the dataset. If certain observations are missing for a particular variable, the relationship between that variable and other variables in the dataset can be used to impute values for the missing observations. In contrast, the task faced by Dr. Neels is to create all values for four key

[6] Dr. Neels is unable to impute values for accountables so he drops this variable from the equation. See Supplemental Report of Kevin Neels on Behalf of United Parcel Service," Docket No. RM2015-7 at 37.

volume variables based upon data which is outside the Form 3999 dataset.[7] This is a herculean task. Regardless of any other problems with his overall approach, if Dr. Neels' imputations are not accurate and robust, then estimation of a national Form 3999 model fails.

Dr. Neels attempts to justify his imputation equations by pointing how that they have moderate to relatively high R^2 statistics, which range from a low of about 30 percent to a high of about 60 percent.[8] It turns out, however, that the R^2 statistic is not applicable to support an inference of accuracy of the imputations. Dr. Neel's imputation exercise is an example of out-of-sample forecasting, because it requires him to use a model estimated on data collected across the 300 ZIP codes in the special study to predict values for the missing volumes for another 10,000 ZIP Codes, included in the national Form 3999 dataset. A moderate (or even high) R^2 in a forecasting model speaks only to how the model "fits" the dataset on which it was estimated, and does not address how well the model forecasts out-of-sample values. To accurately evaluate Dr. Neels' imputation equations, one needs to examine their ability to forecast out of sample.

Because no actual values for in-receptacle parcels, deviation parcels, collection mail, or accountables are included in the Form 3999 dataset, it may, at first blush, seem

[7] Dr. Neels is in the somewhat awkward position of having to critically rely upon the Postal Service's special study data -- which he has roundly criticized -- for performing the imputations. If, as Dr. Neels has argued, the special study data are insufficient for estimating variability equations, it is difficult to see how they are sufficient for performing the required imputations. Dr. Neels does not address data quality in his discussion of his imputation equations.

[8] See, "Report of Kevin Neels on Behalf of United Parcel Service," Docket No. RM2015-7, at 32-36

like there is no way to gauge the out-of-sample forecasting ability of Dr. Neels' imputation equation. And, in fact, the complete omission of these variables from the Form 3999 dataset means that one cannot pursue an out-of-sample test with full rigor. But one can put the imputation equations to a less rigorous, but still illuminating test. Such a test can potentially provide useful insight because if the imputation models fail this less rigorous test, then they will certainly fail the more stringent one.

Specifically, to produce insight about the abilities of these models to forecast out of sample, I estimated one of Dr. Neels' imputation models on half of the special study data (the first week) and then examined how well it forecast the values for volume for the other half of the sample (the second week). Admittedly, this is a "minimum-difficulty" test of the imputation model, but it is a legitimate out-sample test. If there are any volumes the imputation models should be able to accurately forecast, it would be the volumes for the same ZIP Codes for the very next week.

The first step in this examination is the replication of Dr. Neel's negative binomial prediction equation. I performed this exercise for the deviation parcel equation because Dr. Neels states that he is confident of the imputation results for this equation:[9]

> The deviation parcel volumes used in my model are based upon an imputation regression. That regression relies upon a proxy measure of deviation parcel volumes – namely, the DOIS parcel volume counts – that is widely available and that is closely related both conceptually and empirically to deviation parcels. The availability of this measure and its close empirical relationship to the volume measure being estimated gives me confidence in the robustness of the imputation results for deviation parcels, despite the known problems that infect this measure.

[9] See, 'Supplemental Report of Kevin Neels on Behalf of United Parcel Service," Docket No. RM2015-7 at 47

In addition, Dr. Neels argues that the set of deviation parcel predictions are closely linked to Postal Service parcel data and "the volume of national Form 3999 parcels per delivery point is a strong predictor of deviation parcels."[10]

The next table presents the results of Dr. Neels' original estimation of the negative binomial model for deviation parcels (the imputation equation), my replication of that model, and the results of estimating the model on just the first week of special study data. My replication does not exactly match Dr. Neels' model, but that is to be expected, as the negative binomial regression is estimated through a search algorithm, which may vary slightly across different estimation software packages.[11] Nevertheless, the estimated coefficients are extremely close, and the replication was successful. The last column in the table shows that model estimated on just the first week of special study data produces very similar results to Dr. Neel's and, in fact, has a slightly higher pseudo R^2 statistic.[12]

[10] Id. at 32.

[11] Dr. Neels used Stata, and I used SAS to estimate the negative binomial imputation equation.

[12] The complete estimation results for the replication and the model estimated on the first week of data are presented in USPS-RM2015-7/5.

Estimation of the Negative Binomial Model for Deviation Parcels

	Neels Original	Replication	On First Week of Data
Constant	4.147	4.139	4.136
Possible Deliveries	0.0000683	0.0000689	0.0000714
DPS Letter per DP	0.148	0.149	0.141
Cased Flats Per DP	0.022	0.028	0.040
Parcels Per DP	3.900	3.833	4.127
Tuesday	0.063	0.052	0.020
Wednesday	0.266	0.257	0.253
Thursday	0.271	0.268	0.259
Friday	0.247	0.245	0.231
Saturday	0.155	0.149	0.130
Psuedo R2	0.544	0.550	0.566
# of Obs	3,333	3,332	1,674

The next step in the evaluation process is to use the model estimated with the first week of data to predict, “out-of-sample” the second week’s values for deviation parcels. The last step is to then compare those predictions with the actual values.

Application of the imputation model to the second week’s data produced 1,658 individual predictions for deviation parcels that can be analyzed. The next table provides a few examples of those predictions, and their associated prediction errors, for deviation parcels. Prediction errors are defined as the predicted value minus the actual value, so a positive prediction error reflects an over-prediction. This table is illustrative only and not intended to be representative of the overall pattern of errors. More systematic evaluation of the predictions will follow, but this table provides a feel for the nature of the prediction errors. To that end, it includes predictions for low volume, medium volume, and high volume deviation parcel ZIP Code days.

Examples of Prediction and Prediction Errors for Deviation Parcels

Masked ZIP Code	**Actual Value**	**Predicted Value**	**Prediction Error**	**Percentage Prediction Error**
17799	34	175.8	141.8	417.1%
83226	42	754.9	712.9	1697.4%
15912	80	158.0	78.0	97.5%
24366	94	186.1	92.1	98.0%
17007	246	808.8	562.8	228.8%
17444	275	175.9	-99.1	-36.0%
47889	281	145.8	-135.2	-48.1%
99823	354	850.5	496.5	140.3%
97452	369	177.5	-191.5	-51.9%
80548	417	828.8	411.8	98.8%
46107	589	284.1	-304.9	-51.8%
48719	666	1399.9	733.9	110.2%
12629	799	397.3	-401.7	-50.3%
83077	805	1203.8	398.8	49.5%
12629	1049	497.0	-552.0	-52.6%

More systematic evaluation of the quality of the deviation parcel predictions can be provided by computing measures of central tendency of the prediction errors. Because there are both positive and negative prediction errors, the appropriate measures of central tendency should be applied to the absolute value of those errors. For example, a widely used statistic is the mean absolute prediction error (MAE). Its formula is given by the following equation, in which p_i represents the predicted value, a_i represents the actual value, and N is the number predictions made:

$$MAE = \frac{\sum_i^N |p_i - a_i|}{N}$$

The mean absolute error for the out-of-sample deviation parcel forecasts is 113.8 parcels. To assess the size of that average error, it can be compared with the average value for deviation parcels of 326.8. This comparison suggests that the average prediction error is quite large relative to the variable being forecast. Another useful measure of forecast accuracy is the mean absolute percentage prediction error. Its formula is similar to the MAE, but makes use of the absolute percentage error. For deviation parcel predictions, the mean absolute percentage prediction error is 88.9 percent.

However, both the mean absolute error and the mean absolute percentage error are influenced by some very large prediction errors that are well over 100 percent (meaning the prediction error is larger, in absolute terms, than the actual value).[13] To abstract from the impact of those extremely bad forecasts on the measure of central tendency for all forecasts, one can calculate the median absolute and median absolute percentage forecast errors. As expected, the median prediction errors are well less than the mean prediction errors. The median absolute prediction error is 79.6 packages and the median absolute percentage prediction error is 29.5 percent. Although they are smaller than the means, the median values still suggest that Dr. Neels' imputation equations generate very large prediction errors. These results raise serious questions about the utility of the imputations in estimating the street time variability equation.

Perhaps one reason for some of Dr. Neels' large prediction errors is from the fact that he treats the Form 3999 dataset as a pure cross section, when it is not. In fact, as explained in the Postal Services report in this docket, the Form 3999 data come from a number of years, primarily from 2010 through 2013.[14] My understanding is that Dr. Neels is using the special study data from 2014 to estimate the models used to predict the package and accountable volumes for the ZIP Codes covered by the Form 3999 data. Some of those ZIP Codes have Form 3999 data may have come from, say, 2010 or 2011. Given the recent growth in parcel volumes, it is not clear that even an accurate imputation equation based on 2014 data would be able to accurately predict package volumes for 2010 or 2011. This might help explain some of the very large over predictions generated by the model.

[13] There are 270 instances in which the prediction error is larger than the actual value. The existence of these large errors is reason that the average percentage prediction error is much larger than the ratio of the average prediction error (113.8) to the average value (326.8). That ratio is 34.8 percent.

[14] See, Report on the City Carrier Street Time Study, Docket No. RM2015-7 at 10.

Another widely used method of evaluating forecasts examines whether the forecasts are unbiased and efficient. This test is done with the Mincer-Zarnowitz equation. As with the MAE formula, p_j represents a predicted value and aj represents the corresponding actual value:[15]

$$a_j = \beta_0 + \beta_1 p_j + \varepsilon_j$$

An unbiased forecast requires $\beta_0 = 0,\ \beta_1 = 1$. An efficient forecast requires just $\beta_1 = 1$. I estimated the Mincer-Zarnowitz equation on the 1,658 out-of-sample deviation parcel predictions generated by Dr. Neel's model and the next table presents the results of that estimation.

Mincer-Zarnowitz Equation Estimation

Individual t-tests

Coefficient	Estimated Value	t-value
β_0	25.837	3.21
β_1	0.883	43.01

F-Tests for Bias and Efficiency

Test	Test Statistic	Probability Value
Bias	22.73	<.0001
Efficiency	32.67	<.0001

[15] For an example of an application of this equation, See Ericsson, Neil, "How Biased are U.S. Government Forecasts of the Federal Debt," International Journal of Forecasting, in press (2015).

These results indicate that Dr. Neel's imputation equation produced predictions for deviation parcels which are both biased and inefficient. The means that the imputation equation does a poor job of predicting the actual deviation parcel values for the same ZIP Codes on which the model was estimated.

In sum, the tests of the imputations required to estimate the aggregate street time equation using national Form 3999 data show that they are biased and inaccurate. They do not appear to provide a reliable basis for accurately estimating the coefficients in the variability equation. Because accurate predictions are necessary for reliable estimation of the variability equation, the proposed procedure of using the national Form 3999 data for that estimation fails on this basis alone.

But even if the predicted values for the volumes were better, serious econometric problems remain with the proposed procedure. Dr. Neels' approach is known as a "two-step" estimator in which the predictions are generated in the first step and the model, using those generated predictions, is estimated in the second step. As a result, some of the variables on the right-hand-side of Dr. Neels' aggregate street time equation are stochastic. As Dr. Neels alludes to in his supplemental report, the impact of these stochastic explanatory variables is that the resulting standard errors for the variability equation are biased downward.[16] Consequently, inferences about the estimated coefficients are not reliable, as t-statistics will be overstated. But even with biased standard errors, 30 of 66 estimated coefficients (45 percent) in Dr. Neels' street time variability equation are not statistically significant at the traditional 95 percent level.

[16] See, 'Supplemental Report of Kevin Neels on Behalf of United Parcel Service," Docket No. RM2015-7 at 37.

Thus, it is quite likely that more than half of the estimate coefficients are truly not significant.

The fact that so many estimated coefficients are not measurably different from zero also reflects the severe multicollinearity that infects the variability model estimated on the Form 3999 data. Dr. Neel's specification has 66 right-hand-side variables (most of which are linear, quadratic, and cross-product terms) in the various delivered volumes. It is well established that these volumes are highly correlated, so having so many terms almost guarantees a serious multicollinearity problem. Moreover, it appears that the imputation process exacerbates the multicollinearity problem as the imputed in-receptacle volumes and the imputed deviation package volumes have an extremely high correlation coefficient of 94 percent. In addition, as shown in the next table, both imputed package volumes have very high correlations with both DPS volumes and delivery points. Upon reflection, this is not surprising, because both of those variables are important variables in the imputation equations. This result raises the possibility that the imputed package volumes may not be representing the true package volumes, but rather serving as reflections for the important roles that DPS volume and delivery points independently play in determining street time.

Correlation Coefficients and Probability Values For Parcel Imputations With Form 3999 DPS Volume and Delivery Points

	Imputed Deviation Packages	DPS Volume	Delivery Points
Imputed In-Receptacle Package Volume	0.941 <.0001	0.895 <.0001	0.885 <.0001
Imputed Deviation Package Volume		0.890 <.0001	0.896 <.0001
DPS Volume			0.914 <.0001

The problem of understated standard errors caused by the generated regressors also infects Dr. Neels' calculations of standard errors for marginal times, because these latter standard errors are based upon the stochastic structure of the estimated coefficients. This is a concern because the downwardly-biased standard errors for the marginal times for the parcel shapes are already quite large. Even these incorrectly estimated standard errors suggest that the parcel marginal times are not sufficiently well estimated to provide reliable attributable costs. For example, Dr. Neels' estimated standard errors suggest that a 95 percent confidence interval for the marginal time for deviation parcels would range from a low of 1.6 minutes to a high of 3.8 minutes. This is an extremely wide range and the top and bottom figures have very different implications for the attributable street time costs of deviation parcels.

Moreover, a number of the marginal times produced by the model based upon the national Form 3999 data appear to be operationally infeasible. The model produces a marginal time for delivery of an FSS piece of 13 seconds. This is difficult to understand in light of the fact that the time marginal time for a cased piece (which are primarily flats) is under 4 seconds and the marginal time for a sequenced piece (which includes many flats) is just 1.4 seconds.[17] To provide some perspective on this estimate, note that the average gross street time per delivery point in the national Form 3999 dataset is 35.8 seconds. Dr. Neels' result suggests that the delivery of just one FSS piece would consume over one-third of the average street time per delivery point.

In addition, Dr. Neels' Form 3999 model predicts that the marginal time for a deviation package is 162.08 seconds or 2.7 minutes. This seems extremely high. Some deviation package deliveries do require the carrier to move the vehicle and/or to depart from the vehicle, so it is possible that some deviation deliveries could take nearly three minutes. But other deviation package deliveries simply require putting the package in a parcel locker in the cluster box unit or behind the storm door (or on the front stoop) at a customer's residence. These types of deviation package deliveries take a few seconds. This operational reality suggests that a marginal deviation package delivery of nearly three minutes is infeasible.

Note that this measure is not the average time per deviation package but the marginal time, the additional time associated with delivering an additional package. With economics of density in delivery, the marginal delivery time will necessarily be below the average delivery time. Because the Postal Service is a multiproduct firm and

---

[17] Unfortunately, Dr. Neels does not provide any operational explanations for the marginal times produced by the model.

delivery involves common production of more than one good, the traditional average cost measure does not exist. Instead, the multi-product firm analog to an average cost is the average incremental cost. Thus to gain insight into the average time for delivering a type of package, one must calculate the average incremental time for that type.

Fortunately, in a quadratic model, there is a straightforward relationship between a shape's marginal time or cost and its average incremental time or cost. To illustrate that relationship, I present a simple two-variable quadratic cost equation. In such a cost equation, total cost, C, is defined as a function of the two outputs, X_1 and X_2:

$$C = \beta_0 + \beta_1 X_1 + \beta_2 X_1^2 + \beta_3 X_2 + \beta_4 X_2^2 + \beta_5 X_1 X_2$$

The marginal cost for the second output, X_2, is given by:

$$MC(X_2) = \frac{\partial C}{\partial X_2} = \beta_3 + 2\beta_4 X_2 + \beta_5 X_1$$

The incremental cost for X_2 is defined as the difference between the total cost incurred and the cost would be incurred if X_2 was not provided. That latter cost is given by:

$$C(-X_2) = \beta_0 + \beta_1 X_1 + \beta_2 X_1^2$$

The incremental cost of X_2 is thus given by:

$$IC(X_2) = \beta_3 X_2 + \beta_4 X_2^2 + \beta_5 X_1 X_2$$

The average incremental cost (AIC) of X_2 is found by dividing incremental cost for X_2 by its volume:

$$AIC(X_2) = \beta_3 + \beta_4 X_2 + \beta_5 X_1$$

Comparing the marginal cost expression for X_2 with the average incremental cost expression for X_2 produces the relationship between the two cost measures in a quadratic equation:

$$AIC(X_2) = MC(X_2) - \beta_4 X_2$$

One can use this relationship to develop the average incremental times for both deviation packages and in-receptacle packages implied by Dr. Neels' Form 3999 model. The average incremental time for a deviation package is 5.5 minutes and the average incremental time for an in-receptacle package it is 2.6 minutes. These times are extremely high and would imply, for example, that a route with just 50 packages, 30 in-receptacle and 20 deviation, would spend over three hours of street time, about half of average total street time per route, to deliver just those packages.

Further perspective on these estimated times can be provided by recalling that Postal Service operations experts identified ceilings for feasible delivery times for both in-receptacle packages and deviation packages [18] The Postal Service indicated that it

[18] See, Report on the City Carrier Street Time Study, Docket No. RM2015-7 at 101.

is considered infeasible if a ZIP Code's average daily in-receptacle delivery time was over 3 minutes per piece or if its average deviation package delivery time was over 5 minutes per piece. ZIP Code days with values above these values were identified as reflecting incorrect data. Yet Dr. Neels' model is thus suggesting overall average times for package delivery that, nationwide, are at or above the feasible time ceilings.

The challenge faced by Dr. Neels in using the national Form 3999 dataset to estimate city carrier street time variabilities was daunting. To be successful, he had to overcome the omission of data on four key volumes and the existence of extreme multicollinearity. Despite an extensive effort, he was not able to overcome the barriers he faced and was not successful in producing acceptable variability estimates.

Evaluating Dr. Neels Modified Proposal 13 Approach

To his credit, Dr. Neels candidly admits that there are limitations to his estimation based upon the Form 3999 data and imputations, and presents an alternative approach to the Commission for its consideration.[19] This alternative consists of modifying the Postal Service's Proposal 13 analysis by including DOIS parcels in the regular delivery equation, and then estimating it on the data collected for approximately 300 ZIP Codes over a two-week period in 2013. This model was originally presented in Dr. Neel's initial report and I addressed it in my analysis of the report. However, because Dr. Neels raises this model again, it bears a few additional comments.

---

[19] <u>See</u>, 'Supplemental Report of Kevin Neels on Behalf of United Parcel Service," Docket No. RM2015-7 at 43

On a conceptual basis, Dr. Neels argues for this modified approach on the basis that during certain parts of the carrier's day, he or she will be handling different mail streams together:[20]

> More generally, however, I understand that over the course of his or her day a letter carrier handles a number of different mail streams. Parcels make up one of those streams. Delivering different streams of mail is a jointly produced service and the handling of these streams is at various points of time thoroughly intermingled. This fact causes me to question whether it is even possible to isolate and measure the time spent dealing with a single mail stream. There may certainly be instances over the course of the day when the letter carrier is handling only one type of mail. However, I would expect those times to add up to much less than the full day. During some major portion of a carrier's day he or she will be dealing with multiple mail streams. It is reasonable to assume that the time required to carry out these other activities is potentially influenced by all mail volumes, to an extent that can probably only be measured statistically. In my opinion, it is not possible to manually unscramble the eggs in a situation like this in order to figure out how much time you spent cooking the egg whites.

Dr. Neels is correct in his understanding that there are many parts of a carrier's day in which mail streams are handled in common. But it appears that Dr. Neels has less clarity about the implications of this common activity for the determination of product costs. In many, if not most, of the activities a carrier performs on several mail streams, the resulting costs are not attributable to any individual product. Consider a motorized carrier driving from delivery point to delivery point. The time required for driving between delivery points would be the same whether the carrier had one, two, three, or four different mail streams in the vehicle. Alternatively, consider a carrier who is walking and deviates from the route's primary line of travel to take the mail to a

[20] Id. at 6.

customer's receptacle. Suppose that the carrier has a parcel, three DPS letters, and two pieces of cased mail for the customer. It turns out that the time (and thus the cost) of walking to the house is not attributable to any of the individual products the carrier is delivering. To see why, note that if the carrier instead did not have the parcel, but took only the three DPS letters, and the two pieces of cased mail, the time to walk to the house would not change. If the carrier did have the parcel, but none of the cased mail, the time again would be the same. Thus, in this instance there is no causal relationship between the individual product volumes and the incurred time. In Dr. Neel's parlance, there is no need to "unscramble" the eggs, as the yolks and whites are baked simultaneously in a single soufflé.

It is only in regard to those activities that are responsive to volume changes that attributable costs arise. Once the carrier is at the mail receptacle, then the time to load the mail into the receptacle is affected by the volumes delivered. The placement of the parcel into the receptacle is a physically separate activity from the placement of the letters and flats, so if the carrier did not have to deliver the parcel, the time at the stop would be reduced. In this instance there is a causal link between the parcel and the delivery time because the presence of the parcel caused the time to change. Specifically, the additional time required to deliver the package is the <u>marginal</u> time associated with that volume.

The widespread existence of common costs in carrier activities ensures that the simultaneous handling of more than one mail stream by a carrier is not sufficient to justify inclusion of packages in the regular delivery equation. Rather, the key issue is the degree of separability between parcel delivery and letter and flat delivery. To understand the concept of separability, consider a cost function with three outputs, X_1, X_2, and X_3. Separablilty in costs between output one and output three occurs when:

$$\frac{\frac{\partial C}{\partial X_1}}{\partial X_3} = 0.$$

This condition implies that we can investigate the degree of separability for parcels, for example, by looking at the cross-product terms between parcels and the other volume measures in Dr. Neels' aggregate street time equation. Those cross-product terms are the empirical analog of the separability condition presented above. The next table presents the t-statistics from that equation. If those t-statistics indicate the cross-product terms are non-zero, the parcels are not separable from letter and flat volumes in delivery. If the t-tests indicate that the coefficients are zero, then they suggest that that separability holds for parcels. The table presents the t-tests for both deviation parcels and in-receptacle parcels and they indicate that the cross product terms are not significantly different from zero. It is true that one of the cross products for in-receptacle parcels (with FSS mail) has a t-statistic large enough for the coefficient to be considered statistically different from zero, but recall that the standard errors in the Form 3999 aggregate street time equation are understated. Understated standard errors cause the t-statistics to be overstated. Given the closeness of that t-statistic for the cross product between in-receptacle parcels and FSS mail to the critical value of 1.96, it reasonable to infer it would not be that large enough with appropriate standard errors.

T-statistics on Cross Product Volume Terms Involving Parcels From Dr. Neel's Total Street Time Equation

Deviation Parcels

DPS	1.48
FSS	-0.93
Sequenced	1.77
Cased Mail	0.81
Collection Mail	-1.55

In-Receptacle Parcels

DPS	-0.20
FSS	2.37
Sequenced	1.66
Cased Mail	0.89
Collection Mail	-0.44

Source: 'Supplemental Report of Kevin Neels on Behalf of United Parcel Service," Docket No. RM2015-7, Table 11 at 38.

As detailed in the previous section, there are many problems with the aggregate street time equation, so one should not make too much of inferences drawn on that model. Therefore, I do not mean to infer that these results are definitive proof that parcels are separable from all other mail shapes. However, this evidence is consistent with work in previous postal rate cases and other postal research that suggests that the delivery of the parcel mail stream may be quite separable, in a cost sense, from the letter and flat mail streams.[21]

[21] For example, in Docket R90-1, United Parcel Service witness Nelson presented a load time model in which parcels were separable from all other volumes. (See Docket No. R90-1, Direct Testimony of Michael A. Nelson on Behalf of United Parcel Service, UPS-T-1, at 20.) In terms of academic research see, Bradley, Michael D. and Jeff

It is important to recognize that the Postal Service's approach to estimating parcel variabilities is based, in large part, on the practicality of accurately estimating those variabilities.[22] Philosophically, I agree with Dr. Neels that there are certain advantages associated with estimating an overall "top-down" variability equation in which all variabilities and marginal costs are simultaneously estimated. In theory, such an approach could allow for econometrically testing separability. In practice, as Dr. Neel's own research has demonstrated, this can be a very difficult course. Even with 10,000 observations, Dr. Neels could not successfully overcome the problems of missing data and multicollinearity, which often plague top-down approaches.

The Postal Service's approach is not predicated on an assumption that there is absolutely no possible relationship between parcel mail and the other mail streams. Rather, it is based upon the practical reality that that relationship, if it exists, is relatively small and is very difficult to estimate accurately.

Dr. Neels indicated that his main concern is increasing the attribution of costs to parcels, but the Postal Service, and the Commission, must be concerned with the accurate attribution of costs to all products.[23] In practice, it could well be that attempts

---

Colvin, "An Econometric Model of Postal Delivery," in Competition in Postal and Delivery Services: National and International Perspective, M. Crew and P. Kleindorfer, eds. Kluwer Academic Publisher, 1995 which finds an absences of economies of scope between parcels and the other mail streams. More recently, in Docket No. R2005-1, the Postal Service proposed and the Commission accepted separate estimation of deviation package and accountable variabilities.

[22] For example, Dr. Neel's modified Proposal 13 model relies upon DOIS parcel counts which both the Postal Service and UPS agree are subject to serious flaws. Such bad data preclude accurate estimation.

[23] See, 'Supplemental Report of Kevin Neels on Behalf of United Parcel Service," Docket No. RM2015-7 at 43, where he states: "This result provides a straightforward

to estimate the interaction between parcels and the rest of the mail stream could materially reduce the accuracy of the estimates of not only parcel variabilities but the variabilities for other bundles such as DPS, cased, or collection mail. If nothing else, adding parcels to a variability equation could induce sufficient multicollinearity to render inaccurate the separate estimation of all variabilities. In this circumstance, the prudent approach is to estimate the variabilities of parcels and accountables separately form the estimation of variabilities for the other mail streams.

Even if one accepts a theoretical argument for including parcels in the regular delivery equation, accurately estimating the associated variability appears to be a difficult challenge. For example, of the nine terms including parcels in Dr. Neels' modified Proposal 13 equation, six of them (66.7 percent) of them are not statistically significant at the 5 percent level and five of them are not statistically significant (55.5 percent) at the 10 percent level. If one uses just the four coefficients that are statistically significant at the 10 percent level to recompute the variability, it falls to just 1.8 percent. Additionally, I attempted to implement Dr. Neel's modified approach with the national Form 3999 dataset, by using it to estimate a regular delivery equation including parcels. That effort produced a variability of 1.0 percent for parcels when all terms are used in its calculation, and a variability of 3.6 percent when just the terms that were significant at the 10 percent level are used. This is a very wide variation in results. To be clear, I am not recommending the use of any of these models, as they all have econometric issues. Rather, I am suggesting that the wide variation in estimated

---

way to address what I believe to be the most serious shortcoming of the Proposal 13 approach – namely the under-attribution of regular delivery time to parcels."

variabilities suggests that it is not possible, with the current data, to accurately estimate a variability for regular delivery time with respect to parcels.

Dr. Neels raised an interesting research issue with his concern that the Postal Service's approach may have missed attributing some costs to parcels. Previous research on city carriers suggests that it is likely that parcel delivery is separable from letter and flat delivery, and Dr. Neels' own work suggests that if a linkage does exists, it is extremely small. Given the difficulties in reliably estimating this relationship with existing data, no modification of the Postal Service's Proposal 13 is currently warranted. But this could be an appropriate topic for future research.